

1. [Discovering the Chemical Structure of DNA: Goals](#)
2. [Relating DNA Structure to Function: The Challenge of the Tetranucleotide Hypothesis](#)
3. [Chargaff's Test of the Tetranucleotide Hypothesis](#)
4. [Was the Tetranucleotide Hypothesis Supported by Quantitative Data?](#)
5. [Chargaff's Additional Discovery](#)
6. [Chargaff's Complementarity and Watson & Crick's Model of DNA Structure](#)
7. [Watson and Crick's Model Suggests a Mechanism for DNA Replication](#)

Discovering the Chemical Structure of DNA: Goals

Learning Goals & Outcomes

After this unit you should have achieved the following goals (numbered) as measured by your ability to successfully produce the associated products (bulleted):

1) Understand the base-pairing rules, adenine (A) with thymine (T) and cytosine (C) with guanine (G), which govern the molecular organization of DNA.

- [Formulate](#) and [test](#) an hypothesis about DNA's expected nucleotide composition based Levene's tetranucleotide model of DNA structure.
- [Formulate a model to explain the nucleotide composition of DNA discovered by Chargaff.](#)

2) Understand the basic macrostructure of DNA: a single DNA molecule is composed of two individual DNA molecules (strands) each of which is composed of a single file sequence of nucleotides. The two individual strands (DNA molecules) are held together by hydrogen bonds between complimentary base-pairs A with T and C with G.

- [Elaborate Watson and Crick's simple diagram for the structure of DNA.](#)
- [Formulate a model for how a DNA molecule can be replicated based on its structure as proposed by Watson and Crick.](#)

3) Understand that function arises from structure even at the molecular level.

- [Evaluate whether the structure of DNA, as proposed by the tetranucleotide hypothesis, could account for the proposed function of DNA as the molecule of inheritance.](#)

- Formulate and test an hypothesis about DNA's expected nucleotide composition based Levene's tetranucleotide model of DNA structure.
- Explain the importance of Watson and Crick's observation that their proposed model of DNA macrostructure immediately suggests a mechanism for copying of DNA.
- Formulate a model to explain how the DNA molecule can be replicated based on its structure as proposed by Watson and Crick.

4) Wrestle with interpreting variation in scientific data.

- Interpret original, variable scientific data and draw conclusions.
- Compare your conclusions to those drawn by the original scientists themselves.

Relating DNA Structure to Function: The Challenge of the Tetranucleotide Hypothesis

These days just about everyone knows DNA is the molecule of heredity. Even four year olds grasp the idea that DNA is some how responsible for their physical similarity to others!

Naturally there was a time when DNA's function was completely mysterious. In fact, for nearly eighty years after Friedrich Miescher's serendipitous discovery of 'nuclein' in 1869 (he was actually studying proteins), DNA's biochemical structure and its biological job remained unresolved and contentious.

This prolonged period of uncertainty and controversy grew in part from the need to unite constantly evolving theory and knowledge from three disparate and largely independent scientific fields – biochemistry, cellular biology and genetics (Dahm, 2005). Within this context, one of the great challenges was to convincingly reconcile the hypothesized biochemical structure of DNA, as revealed by biochemists, with its proposed function as the molecule of inheritance derived from the work of cellular biologists and geneticists.

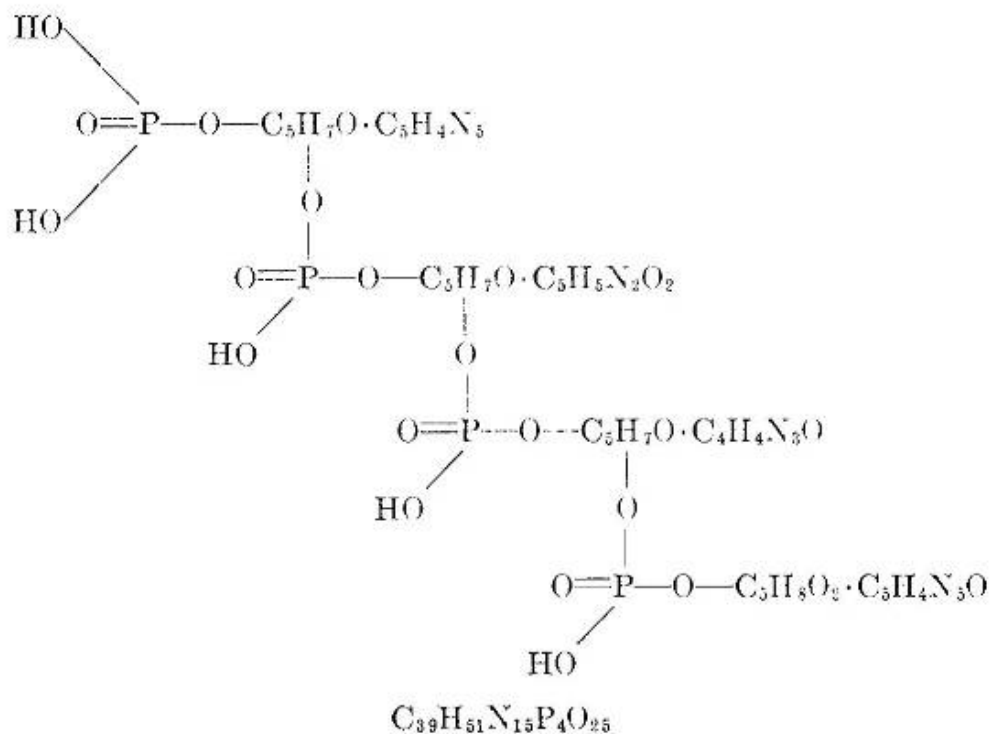
Specifically, to work as the molecule of inheritance, DNA's chemical structure had to be capable of encoding the enormous number of genetic recipes responsible for the huge diversity of life on Earth. In other words, the structure of DNA had to enable its function much like our letter-rich alphabet permits a diversity of written words. What kind of vocabulary could we have if the alphabet consisted of a single letter? Analogously, one very important question of the time was did DNA have sufficient letters and organizational flexibility to spell out a huge variety of genetic codes?

To investigate this, let's return to the first half of the 20th century. By 1929, the following was known about the biochemical composition of DNA (Chargaff, 1971; [Figure 1](#)):

- DNA was rich in the element phosphorous, specifically phosphate, distinguishing it from all other classes of biological molecule (i.e. proteins, carbohydrates, and lipids).

- DNA was composed of deoxyribose sugar molecules.
- DNA was composed of four different nitrogen-containing bases, adenine (A), cytosine (C), guanine (G), and thymine (T).
- Together one nitrogen-containing base, one deoxyribose sugar and one phosphate formed a nucleotide, the monomer from which the polymer DNA was formed. Illustrations of single nucleotides can be found [here](#).
- Each nucleotide was chemically bound to the next via a phosphate.

Additionally in 1912, the biochemist Phoebus Levene had proposed a model for the structural arrangement of DNA's four types of nucleotides - the **tetranucleotide hypothesis**. According to this hypothesis, a DNA polymer was composed of a serially repeated, four nucleotide unit the tetranucleotide or (ATCG)_n where n represents the number of tetranucleotides in the polymer (Chargaff, 1971; [Figure 1](#))



Levene never published a figure illustrating the proposed tetranucleotide. This figure describing the structure of four bonded nucleotides and taken from a 1929 publication can be used to imagine a tetranucleotide-like unit. Levene and London, 1929. On the structure

of thymonucleic acid. *The Journal of Biological Chemistry*. 83:793-802.

Despite limited supporting evidence, Levene's model of DNA structure held great sway for several decades even appearing in textbooks as one of the few known facts about DNA (Hotchkiss, 1979). Its prominence stemmed in part from Levene's standing as an influential scientist in the field of nucleic acid chemistry as well as the lack of methods to quantitatively analyze DNA's nucleotide content (Hotchkiss, 1979).

Consider Levene's tetranucleotide model for the structure of DNA and answer the questions that follow.

1. Illustrate your understanding of Levene's tetranucleotide hypothesis by drawing a molecule of DNA formed of 5 tetranucleotides. Please be sure to specifically include the individual nucleotides A, T, C and G.
2. Recall the issue confronting scientists in the early to mid-1900's: if DNA was the molecule of inheritance as hypothesized by cell biologists and geneticists then the structure of DNA had to be sufficiently variable to encode the genetic diversity responsible for species diversity.

Examine your illustration (question 1). Do you think a DNA polymer built of tetranucleotides, even if it were formed of a 1000 or 1 million tetranucleotides, would be variable enough to account for the differences among species? Why or why not? Please explain.

Works Cited

- Chargaff, E. 1971. Preface to a grammar of biology. A hundred years of nucleic acid research. *Science*. 172:637-642.
- Dahm, R. 2005. Friedrich Miescher and the discovery of DNA. *Developmental Biology*. 278:274-288.
- Hotchkiss, R.D. 1979. The identification of nucleic acids as genetic determinants. *Annals New York Academy of Sciences*. 325:321-342.
- Levene, P.A. and E.S. London. 1929. On the structure of thymonucleic acid. *The Journal of Biological Chemistry*. 83:793-802.

Chargaff's Test of the Tetranucleotide Hypothesis

Scientists of the early 1900's recognized that if the [tetranucleotide model](#) was correct, then DNA could not differ sufficiently among species to explain the incredible diversity of heritable differences. From the biochemical perspective, proteins were a superior candidate because they were known to be both incredibly diverse and physically associated with DNA in the nucleus.

From the cell biology and genetics perspective, however, evidence continued to accumulate that DNA was responsible for heredity. One particularly significant discovery, published in 1944 by Oswald T. Avery, Colin MacLeod, and Maclyn McCarty, provided very strong evidence that DNA was responsible for inducing heritable changes in a species of bacteria (Avery et al., 1944).

While momentous in hindsight, Avery et al.'s discovery was not particularly illuminating to the generally protein-focused biochemistry community. The biochemist Erwin Chargaff, however, reacted differently. As he later wrote, Avery et al.'s results

...almost abruptly, appeared to foreshadow a chemistry of heredity and, moreover, made probable the nucleic acid character of the gene. It certainly made an impression on a few, not many, but probably on nobody a more profound one than on me. For I saw before me in dark contours the beginning of a grammar of biology. (Chargaff, 1971, p.639)

The intellectual impact of Avery et al.'s work was such that Chargaff abandoned all prior research (or completed it quickly) and refocused the efforts of his entire lab on the question of DNA. As he wrote in 1971,

I started from the conviction that, if different DNA species exhibited different biological activities, there should also exist chemically

demonstrable differences between deoxyribonucleic acids. From the very beginning I drew an analogy to the proteins in assuming that the biological activity of the nucleic acid probably rested on the sequence specificity of its constituents - on the order in which the four different nucleotides were arranged in the macromolecule - rather than on the occurrence of new, as yet unrecognized constituents. (Chargaff, 1971, p.639)

Chargaff's nearly 'insurmountable' challenge was to develop an analytical technique to precisely and reliably quantify the nucleotide composition of DNA. Such a tool would allow him to test the idea that DNA of different species differed in nucleotide composition. Such a result would be consistent with Chargaff's proposition, inspired by Avery et al.'s work, that DNA is biologically active, its biological activity differs among species (in that it is responsible for their different heritable phenotypes) and that its activity is conferred by the particular sequence of nucleotides.

Let's consider what insights the ability to quantify the relative amounts of A, T, G and C provide in relation to both the tetranucleotide model and Chargaff's proposition above.

1. Reconsider the [tetranucleotide model](#).

a. If DNA is composed of serially repeated tetranucleotides and you precisely measure the relative quantities of A, T, C and G in an organism's DNA, what do you expect to find? Why? Please explain.

b. If DNA is composed of tetranucleotides and you compare the relative quantities of A, T, C and G (for example, the ratios of A to G or T to C) among species what would you expect to find? Why? Please explain.

2. Now consider Chargaff's hypothesis. If DNA is biologically active, its biological activity differs among species and its activity stems from the particular sequence of nucleotides, what would you expect to find if you

precisely measure the relative quantities of A, T, C and G in a variety of species and compare them? Why? Please explain.

Works Cited

- Avery, O.T., C.M. MacLeod, and M. McCarty. 1944. Studies on the chemical nature of the substance inducing transformation of Pneumococcal types. *The Journal of Experimental Medicine*. 79: 137-159.
- Chargaff, E. 1971. Preface to a grammar of biology. A hundred years of nucleic acid research. *Science*. 172:637-642.

Was the Tetranucleotide Hypothesis Supported by Quantitative Data?

In order to test his [expectation](#) that the DNA of different species quantitatively differed in nucleotide composition, Chargaff's lab spent two years, 1946-48, developing a chromatographic technique to reliably and precisely quantify the amounts of A, T, C and G that had been chemically liberated (via hydrolysis) from polymers of DNA (Chargaff, 1971).

So what did they find? The tables below (Figure 1) show Chargaff's original results (1950) describing the relative quantities of A, T, C and G in DNA extracted from four different species – the ox, human, yeast and a bacterium (avian tubercle bacilli).

Original data describing the relative quantities of adenine, guanine, cytosine and thymine and total amount of these nitrogenous constituent recovered (recovery) in moles per mole of phosphorous for four different species 1.1) ox 1.2) human 1.3) yeast and a bacterium (avian tubercle bacilli). Note that for both ox and human DNA from three different types of tissue has been characterized. Each molar value represents the average of several hydrolysis experiments on DNA from the species (Chargaff, 1950).

Table II ¹

Composition of deoxyribonucleic acid of ox (in moles of nitrogenous constituent per mole of P).

Constituent	Thymus			Spleen		Liver
	Prep. 1	Prep. 2	Prep. 3	Prep. 1	Prep. 2	
Adenine . .	0.26	0.28	0.30	0.25	0.26	0.26
Guanine . .	0.21	0.24	0.22	0.20	0.21	0.20
Cytosine . .	0.16	0.18	0.17	0.15	0.17	
Thymine . .	0.25	0.24	0.25	0.24	0.24	
Recovery . .	0.88	0.94	0.94	0.84	0.88	

Table III²

Composition of desoxypentose nucleic acid of man (in moles of nitrogenous constituent per mole of P).

Constituent	Sperm		Thymus	Liver	
	Prep. 1	Prep. 2		Normal	Carcinoma
Adenine . . .	0.29	0.27	0.28	0.27	0.27
Guanine . . .	0.18	0.17	0.19	0.19	0.18
Cytosine . . .	0.18	0.18	0.16		0.15
Thymine . . .	0.31	0.30	0.28		0.27
Recovery . . .	0.96	0.92	0.91		0.87

Table IV⁵

Composition of two microbial desoxyribonucleic acids.

Constituent	Yeast		Avian tubercle bacilli
	Prep. 1	Prep. 2	
Adenine	0.24	0.30	0.12
Guanine	0.14	0.18	0.28
Cytosine	0.13	0.15	0.26
Thymine	0.25	0.29	0.11
Recovery	0.76	0.92	0.77

Review the results above on a species by species basis and use this to respond to the questions below.

1. Review your expectations under the tetranucleotide model as expressed in your response to [question 1a](#) in the previous module. Would you interpret Chargaff's results (Figure 1 above) as supporting or refuting the tetranucleotide model? Please explain using data to support your conclusions.

2. Levene based his tetranucleotide model on studies of DNA extracted from the thymus cells of cows from which, using very harsh chemical techniques, he recovered approximately equal quantities of A, T, C and G.

- Do Chargaff's results appear to be consistent with Levene's? Yes or no?
- If in part a you concluded Chargaff and Levene's results differed, offer an explanation for this discrepancy.

c. Please explain one way you could resolve the discrepancy between Levene and Chargaff's results.

Chargaff also compared the relative quantities of adenine and guanine (and thymine and cytosine) among these four different species of organisms. His data appear in Figure 2.

Table V
Molar proportions of purines and pyrimidines in desoxypentose nucleic acids from different species.

Species	Number of different organs	Number of different preparations	Adenine/Guanine			Thymine/Cytosine		
			Number of hydrolyses ^a	Mean ratio	Standard error	Number of hydrolyses ^a	Mean ratio	Standard error
Ox ¹	3	7	20	1.29	0.013	6	1.43	0.03
Man ²	2	3	6	1.56	0.008	5	1.75	0.03
Yeast.	1	2	3	1.72	0.02	2	1.9	
Avian tubercles bacillus	1	1	2	0.4		1	0.4	

Relative quantities (ratios) of adenine to guanine and thymine to cytosine in ox, human, yeast and a bacterium (avian tubercles bacillus). Number of hydrolyses refers to the number of times DNA molecules were broken down into their constituent nucleotides to determine quantities of each nucleotide. Mean ratio describes the average ratio for the indicated number of hydrolyses (Chargaff, 1950).

Review the results depicted in Figure 2 and answer the questions below.

3. Review your expectations under the tetranucleotide model as expressed in your response to [question 1b](#) in the previous module. Would you interpret Chargaff's results, as shown in Figure 2, as supporting or refuting the tetranucleotide model? Please explain using data to support your conclusions.

4. Do you agree with Chargaff's unequivocal conclusion based on the data presented above that "*The results serve to disprove the tetranucleotide hypothesis.*" (Chargaff, 1950 p. 206)? Why or why not? Please explain.

5. Review your response to [question 2](#) in the previous module. What do you think Chargaff's results, as shown in Figure 2 above, suggest about the potential for the composition of DNA to vary among species and thus

possibly have biological activity that is based on differences in nucleotide sequence? Please explain.

6. Do you agree with Chargaff's additional conclusions (quoted below) based largely on the data above? Why or why not? Please explain.

Generalizations in science are both necessary and hazardous; they carry a semblance of finality which conceals their essentially provisional character; they drive forward, as they retard; they add, but they also take away. Keeping in mind all these reservations, we arrive at the following conclusions. The desoxypentose nucleic acids [DNA] from animal and microbial cells contain varying proportions of the same four nitrogenous constituents, namely, adenine, guanine, cytosine, thymine. Their composition appears to be characteristic of the species....The presumption, therefore, is that there exists an enormous number of structurally different nucleic acids.....and I think there will be no objection to the statement that, as far as chemical properties go, they [DNA] could very well serve as one of the agents, or possibly as the agent, concerned with the transmission of heritable properties. (Chargaff, 1950, p.208)

Works Cited

- Chargaff, E. 1950. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia*. 6:201-240.
- Chargaff, E. 1971. Preface to a grammar of biology. A hundred years of nucleic acid research. *Science*. 172:637-642.

Chargaff's Additional Discovery

In addition to [discrediting the tetranucleotide hypothesis](#) and reviving the possibility that [DNA maybe sufficiently structurally complex \(and variable among species\) to function as the molecule of heredity](#), Chargaff's careful analytical work revealed one more pattern that would eventually prove very useful for describing the three dimensional structure of the DNA molecule. As Chargaff noted in his 1950 paper,

It is, however, noteworthy – whether this is more than accidental, cannot yet be said – that in all desoxypentose nucleic acids [DNA] examined thus far the molar ratios of total purines [A and G] to total pyrimidines [C and T] , and also of adenine to thymine and of guanine to cytosine, were not far from 1. (Chargaff, 1950, p.206)

1. In your own words, please explain what Chargaff observed as described above in his own words.
2. Review the data in [Figure 1](#). Do agree with Chargaff's interpretation of his data as expressed in the excerpt above? Why or why not? Please explain and support your conclusion with data.
3. Ultimately, Chargaff's interpretation of the data proved correct; the nucleotides A and T (and C and G) occur in equal quantities in DNA regardless of the species from which the DNA comes. Please offer a model of the physical arrangement of these molecules that could explain this recurrent pattern.

Works Cited

- Chargaff, E. 1950. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia*. 6:201-240.

Chargaff's Complimentarity and Watson & Crick's Model of DNA Structure

On the basis of his [results](#) suggesting that the nucleotides adenine and thymine occur in equal quantities in DNA (and guanine and cytosine) regardless of the species from which the DNA comes, Chargaff proposed that nucleotides are organized within the DNA molecule such that A is paired with T and C with G. Chargaff originally called this arrangement “complimentarity” but the term was later modified to “base pairing” (Chargaff, 1971).

Although a seemingly simple observation, this result played a central role in the physical model Watson and Crick proposed in 1953 to describe the macromolecular structure of DNA, that is, its three dimensional organization ([Watson and Crick, 1953](#)).

Together with Rosalind Franklin and Maurice Wilkin’s [x-ray diffraction images](#) of crystallized DNA, Chargaff’s contributions allowed Crick and Watson to imagine DNA as a single molecule composed of two helical strands, each strand itself a single molecule of DNA built of nucleotides arranged single file, held together by hydrogen bonds between complimentary base pairs on opposing strands. A’s on one strand hydrogen bonded to T’s on the other, C’s to G’s and vice versa.

As Watson and Crick (1953) wrote,

...if an adenine forms one member of a pair on either chain, then on these assumptions, the other member must be thymine; similarly for cytosine and guanine...It has been found experimentally that the ratio of the amounts of adenine and thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.
([Watson and Crick, 1953](#), p.737)

The figure from Watson and Crick’s seminal paper, illustrating the proposed structure of DNA, appears in Figure 1.1.



This figure is purely diagrammatic. The two ribbons symbolize the two phosphate-sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

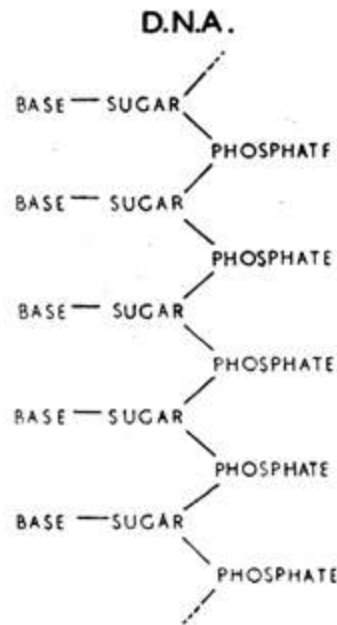


Fig. 1. Chemical formula of a single chain of deoxyribonucleic acid

Original figures illustrating the 1) proposed macromolecular structure of DNA from Watson and Crick's 1953 paper: A structure for deoxyribose nucleic acid. *Nature*. 171:737-738 and 2) proposed structure of one strand of a two-stranded DNA molecule from Watson and Crick's follow-up paper proposing a mechanism for DNA replication: [Genetical implications of the structure of deoxyribonucleic acid](#). *Nature*. 171:964-967.

1. Review Watson and Crick's written and diagrammatic descriptions of DNA. Redraw and modify figure 1.1 to specifically identify the bases A (adenine), T (thymine), C (cytosine) and G (guanine).

2. Compare your illustration to the one found [here](#) and correct it accordingly.

Works Cited

- Chargaff, E. 1971. Preface to a grammar of biology. A hundred years of nucleic acid research. *Science*. 172:637-642.
- Watson, J.D. and F.H.C. Crick. 1953. A structure for deoxyribose nucleic acid. *Nature*. 171:737-738.
- Watson, J.D. and F.H.C. Crick. 1953. Genetical implications of the structure of deoxyribonucleic acid. *Nature*. 171:964-967.

Watson and Crick's Model Suggests a Mechanism for DNA Replication

Watson and Crick's [hypothetical macromolecular structure](#) for DNA also had significant mechanistic implications for solidifying DNA's role in heredity. As [Watson and Crick](#) (1953, p. 737) wrote, "*It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for genetic material.*"

Consider this final point.

1. Watson and Crick have just proposed a hypothetical structure that, as they write, '*rests mainly though not entirely on published experimental data and stereochemical arguments*' ([Watson and Crick](#), 1953, p.737). Why is it so important that this structure suggests a mechanism for copying the DNA molecule? Please explain.
2. Take a look at the [structure of this molecule](#). Offer a hypothetical mechanism by which a DNA molecule might be copied (replicated).
3. Describe a way you could experimentally test your model. As you do this make sure to include a) an hypothesis statement based on your model b) what results you expect if your model is correct and c) what results you expect if your model is incorrect.

Works Cited

- Watson, J.D. and F.H.C. Crick. 1953. A structure for deoxyribose nucleic acid. *Nature*. 171:737-738.